

Neal Grantham

Data Science Generalist in San Francisco, CA
neal@nsgrantham.com (408) 717-0825

nsgrantham.com

[linkedin.com/in/nsgrantham](https://www.linkedin.com/in/nsgrantham)

github.com/nsgrantham

twitter.com/nsgrantham

Education

Ph.D., Statistics, North Carolina State University, Raleigh, NC Aug 2017

Dissertation: *Statistical Methods for High-Dimensional, Spatially-Distributed Microbiome Data from Next Generation Sequencing*

Master, Statistics, North Carolina State University, Raleigh, NC Aug 2014

B.S. Mathematics, B.S. Statistics, California Polytechnic State University, San Luis Obispo, CA Jun 2012

Senior Project: *Analyzing Multiple Independent Spatial Point Processes* GPA 3.84 *Magna Cum Laude*

Experience

Data Scientist, Phylagen, San Francisco, CA Oct 2017 – Feb 2020

- Led research and development on a novel machine learning platform for geoanalysis of microbiome data from environmental sources.
- Built a production-level, end-to-end data analysis pipeline with steps for data acquisition, ETL, large DNA sequence file processing, REST API queries, machine learning analysis, and report generation.
- Performed custom bioinformatic analysis and data visualizations for 25+ customers, including 10 Fortune 500 companies; prepared 50+ presentations and reports with a strong emphasis on clarity and visual design
- “Wearer of many hats” at 20 employee biotech company, contributing to work with government agencies, industry partners, and VC investors, leading to a successful Series A fundraising round of \$14 million.

Graduate Research Assistant, North Carolina State University, Raleigh, NC Aug 2012 – Aug 2017

- Constructed a spatial algorithm to identify the geographic origins of dust samples using microbial DNA sequences – a “biogeographical fingerprint,” funded by the US Department of Defense.
- Developed a Bayesian hierarchical mixed-effects model for the analysis of high-dimensional microbiome data as a response variable from designed experiments.
- Adapted core statistics curriculum to 90+ students with backgrounds in business, biology, and the social sciences. Trained students in statistical thinking via “flipped classroom” approach.

Crew State Monitoring Intern, NASA Langley Research Center, Hampton, VA Jun 2015 – Aug 2015

- Classified cognitive state of airline pilots using eye-tracking, biophysical, and electrical brain signals during critical flight scenarios using hidden Markov models.
- Worked with aerospace, mechanical, and biomedical engineers on open problems in aviation safety.

Undergraduate Research Assistant, Cal Poly, San Luis Obispo, CA Jun 2011 – Jun 2012

- Characterized the clustering patterns of colonial marine invertebrates in capstone Senior Project, requiring independent study of spatial statistics.

Skills

Python	Pandas, Numpy, Scipy, Keras, Scikit-learn, Flask, proficient writing Python packages
R	Tidyverse, ggplot2, Shiny, proficient writing R packages and contributing to open source
Julia	Bayesian MCMC with Mamba.jl, proficient writing Julia packages
SQL	PostgreSQL, SQLite, and database toolkits SQLAlchemy and dbplyr
Unix	Bash scripts, cron jobs, tmux, vim, zsh, dotfiles
Statistics	Bayesian hierarchical models with MCMC, generalized linear models, experimental design
ML/AI	Deep neural networks, random forests, hidden Markov models
Web Dev	Knowledgeable in HTML, CSS, and JavaScript including React, Gatsby, and D3
Notebooks	Jupyter, RMarkdown
Other	AWS EC2, RDS, S3, Docker, Git, GitHub, GSuite, Airtable, Notion

Projects

nsgrantham.com	My website, made with JavaScript and the React & Gatsby frameworks
tidytuesday.rocks	Interactive catalogue of 3,000+ data visualizations, made with R Shiny
github.com/nsgrantham/ggdark	Dark mode extension for ggplot2
github.com/nsgrantham/uspops	R package of population data scraped from US Census Bureau

Publications

- Grantham NS, Reich BJ, Laber EB, Pacifici K, Dunn RR, Fierer N, Gebert M, Alwood JS, Faith SA. (TBD) *Global forensic geolocation with deep neural networks*. In review.
- Grantham NS, Guan Y, Reich BJ, Borer ET, Gross K. (2019) *MIMIX: a Bayesian Mixed-Effects Model for Microbiome Data from Designed Experiments*. JASA.
- Grantham, NS, Reich, BJ, Liu, Y, & Chang, HH. (2018) *Spatial regression with an informatively missing covariate: Application to mapping fine particulate matter*. Environmetrics.
- Grantham NS, Reich BJ, Pacifici K, Laber EB, Menninger HL, Henley JB, Barberán A, Leff JW, Fierer N, Dunn RR. (2015) *Fungi Identify the Geographic Origin of Dust Samples*. PLOS ONE.

Awards

2019 RStudio Shiny Contest Runner-up: Top 10 of 136 submissions
2017 National Institute of Justice Crime Forecasting Winner
2017 JSM 2017 Distinguished Student Paper Award
2016 ENAR 2017 Distinguished Student Paper Award
2015 NASA Internship: nasa.gov/langley/student-story-neal-grantham
2014 NC State Statistics Department Best Written Preliminary Exam
2014 Paige Plagge Graduate Award for Citizenship
2012 Cal Poly Science & Mathematics Senior Recognition Award
2009 George H. McMeen Mathematics Scholarship
2008 Volmar A. & Viola I. Folsom Mathematics Scholarship

Hobbies & Interests

Long Distance Running
Former Radio DJ at WKNC 88.1
Open-Source Software
Data Visualization